IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 31 August 2023, accepted 12 September 2023, date of publication 20 September 2023, date of current version 27 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3317506

RESEARCH ARTICLE

Faster-PestNet: A Lightweight Deep Learning Framework for Crop Pest Detection and Classification

FAROOQ ALI[1], **HUMA QAYYUM**[1], **AND MUHAMMAD JAVED IQBAL**[2]
[1]Department of Software Engineering, University of Engineering and Technology at Taxila, Taxila 47050, Pakistan
[2]Department of Computer Science, University of Engineering and Technology at Taxila, Taxila 47050, Pakistan

Corresponding author: Farooq Ali (Farooq.ali@uettaxila.edu.pk)

ABSTRACT One of the most significant risks impacting crops is pests, which substantially decrease food production. Further, prompt and precise recognition of pests can help harvesters save damage and enhance the quality of crops by enabling them to take appropriate preventive action. The apparent resemblance between numerous kinds of pests makes examination laborious and takes time. The limitations of physical pest inspection are required to be addressed, and a novel deep-learning approach called the Faster-PestNet is proposed in this work. Descriptively, an improved Faster-RCNN approach is designed using the MobileNet as its base network and tuned on the pest samples to recognize the crop pests of various categories and given the name of Fatser-PestNet. Initially, the MobileNet is employed for extracting a distinctive set of sample attributes, later recognized by the 2-step locator of the improved Faster-RCNN model. We have accomplished a huge experimentation analysis over a complicated data sample named the IP102 and acquired an accuracy of 82.43%. Further, a local crops dataset is also collected and tested on the trained Faster-PestNet approach to show the generalization capacity of the proposed model. We have confirmed through analysis that the presented work can tackle numerous sample distortions like noise, blurring, light variations, and size alterations and can accurately locate the pest along with the associated class label on the leaf of numerous types and sizes. Both visual and stated performance values confirm the effectiveness of our model.

INDEX TERMS Classification of pest, deep learning, faster RCNN, pest detection.

I. INTRODUCTION

Many economies rely heavily on agriculture to support their growth and maintain a high living level. Any nation's agro-food sector is essential since it helps to raise the caliber of exports of agricultural products. In emerging nations, the impact of export revenues and domestic market demand is the main driver of the increase in the food processing conversion rate [1]. One of the biggest problems in agriculture today is pest infestation, which results in poor crop quality. Pests, bacteria, and weeds cause substantial crop losses, resulting in sluggish product markets. It is crucial to address pest infestations that affect crop growth. Pests are the primary cause of declines in crop quality, leading to decreased plant productivity. It maintains crop quality and safety in agriculture, and monitoring and evaluating pest damage is essential. As a result, farmers use a range of pesticides to improve crop quality and storage time. Continued use of these pesticides can cause environmental contamination as well as potentially fatal diseases like cancer, severe respiratory and hereditary disorders, and lethal mortality. Advanced technical technologies in agriculture are urgently required to detect pests in their early stages and prevent the widespread application of harmful pesticides [2]. These pests live on the sap from different parts of plants to spread the sooty mold disease. Farmers rely on their expertise and knowledge to identify pest infestations as they arise. Insecticide spraying is regarded as a quick-acting, scalable pest management strategy due to

The associate editor coordinating the review of this manuscript and approving it for publication was Jeon Gwanggil.

a lack of expertise. However, the usage of fewer pesticides should be encouraged due to mounting environmental and health concerns [3], [4].

Spot application is a well-known method that can cut the cost of applying insecticides by up to 90%, reducing pollution and lessening the harmful impacts on beneficial pests like bees. Such applications involve applying various Computer Vision (CV) approaches with the help of image processing to develop recognition systems. CV is increasingly being utilized for various tasks, including soil and crop monitoring, fruit sorting, plant disease detection, and pest identification. Identification and classification of pests are mandatory before spraying on specific areas [5]. Pest detection has traditionally relied on labor and error-intensive manual approaches. Pest and disease identification is essential to acquiring crop growth and health data. The recent developments in computer vision have provided much aid in precision agriculture [6]. Target identification at various stages of agricultural growth is essential to forecast future yields, turn on astute spraying systems, and manage independent insecticide spraying robots in vast farms and gardens. However, with technological advancements, pests can now be found via picture processing. People are increasingly interested in precision agriculture to address these issues [7]. Global positioning systems (GPS) for tractor navigation, robotics, remote sensing, data analytics, drones, and land vehicles are some of these technologies [8]. The foundation of precision agriculture is accurate pest detection. For pest detection and spot spraying, computer vision must be used to capture and analyze visual data [9].

Initially, conventional Machine Learning (ML) approaches with hand-coded features are employed to diagnose numerous crop pests. It consists of four major steps: detection, segmentation, feature extraction, and classification, and is carried out utilizing computer vision-based quality inspection. However, it is challenging to recognize targets with such approaches with adequate precision due to shape similarities, complex backgrounds, target overlap due to high-density distribution, light variations in the large landscape of orchards, and many other issues [10], [11]. Deep Neural Networks (DNNs) are frequently utilized in computer vision applications to plot complicated relationships and perform automatic feature extraction. Deep artificial neural networks can now be trained more quickly and effectively with improved graphics processing units (GPUs). DNN provides meaningful results for classifying objects. Such approaches utilize the idea of transfer learning in which numerous pre-learned deep learning (DL) models are employed to execute a new task. DL methods utilize various Convolutional Neural Networks (CNNs) [12] for extracting effective dataset attributes that do not need domain skills [13]. DL architectures are widely utilized to address complex issues in a sufficient period due to the substantial advancement of computing equipment [14]. DL-based strategies have proven to be precise and have been successfully modified to carry out various farming activities. As improvements in DL techniques have shown encouraging outcomes in the field of recognition of objects, a substantial study has focused on suggesting more complex target localization structures for enhanced identification competence, such as Super-FAN [10] and unsupervised multi-stage key points learning [11] etc. Additionally, several CNN-based methods, including GoogLeNet [15], AlexNet [16], VGG [17], and residual models [18], are tested for pest identification and classification. Some of the latest object recognition approaches like R-CNN [19], Fast Region-based Convolutional Network Method (Fast R-CNN) [20], Faster Region-based Convolutional Network Method (Faster R-CNNs) [21], and You Only Look Once (YOLO) [22] have also exhibited improved results in various domains of agriculture.

The DL object recognition methods listed above have shown outstanding results in designing generic target tracking systems, but they remain limited to a few practical uses for Pest monitoring. Pest diagnosis has distinctive attributes compared to current object spotting and categorization jobs [21].

In real field photos, pests are typically little objects that are accompanied by complicated environments; as a result, any recognition system can be readily misled by the surroundings when estimating key points. Additionally, there is a significant variety in pest mass and positions due to the various vantage points and distances at which they are taken in the agricultural setting, which makes correct diagnosis more difficult. Additionally, diverse pest types frequently share plenty of physical characteristics, and identical types can appear in various forms, including larvae, eggs, pupae, and adults, demonstrating significant within and inter-group differences. In addition, the computerized recognition procedure is made more difficult by the presence of poor illumination and adverse conditions. In order to increase the effectiveness regarding categorization consistency and computing expense, a more reliable and proficient automated approach for exact pest identification in the environment is still needed. The proposed strategy's objective is to overcome existing structures' problems. We have accomplished this objective by introducing an improved DL model named the Faster-PestNet. Descriptively, a customized Faster-RCNN framework is suggested by using the MobileNet as its base network and tuned on the pest samples to recognize the crop pests of various categories and given the name of Faster-PestNet. First, the MobileNet is applied for calculating a distinctive set of image features later recognized by the 2-step locator of the improved Faster-RCNN model. We have proved the robustness of our approach by executing a huge experimental analysis over complex pest samples.

Following are the significant contributions of our work.

- An accurate method capable of computing reliable image features to enhance the Pest's classification performance is proposed.
- Presented such a technique that is capable of accurately detecting and classifying the multiclass Pests due to the high recall power of the proposed technique.

- A vast evaluation of the proposed work is performed on a challenging dataset named IP102 and confirmed through experimentation that the suggested method is able to detect and classify numerous Pests with the presence of several image distortions like noise, blurring, color, and light variations, etc.
- A local crops dataset was collected, and the performance analysis of the presented work was executed. It captured diverse environmental settings and proved the generalization ability of our approach to agricultural applications.

The remaining paper is organized into the following sections: the existing works are discussed in Section II, and the proposed work in Section III. The results and conclusion are demonstrated in Sections IV and V, respectively.

II. LITERATURE REVIEW

With the quick development of Artificial Intelligence (AI) and to address the limitations of ML, CNNs have been fruitfully employed in agricultural research. CNN models outperform conventional techniques when automatically identifying and categorizing pest infestations [23]. Identifying and determining certain pests in realistically obtained images constitute pest categorization. Computer vision issues are better handled by combining CNNs with ensemble models as feature extractors. The techniques utilized include Single-shot multi-box detectors (SSD) [24], YOLO, Faster R-CNNs, Region-based Convolutional Neural Networks (R-CNNs), and Faster R-CNNs. Using it in object detection and recognition has been effective.

Several scholars have conducted recent studies on object detection methods for pest detection. Setiawan et al. [25] performed training on a CNN algorithm for pest detection and used the IP102 dataset as a baseline. This study employed dynamic learning rate, freezing layers, and sparse regularization in conjunction with CutMix augmentation to optimize small MobileNetV2 models. The maximum accuracy, 71.32%, was achieved by amalgamating those procedures throughout training. Nanni et al. [26] utilized the IP102 and a small dataset to spot and identify pest images. The author utilized CNN methods AlexNet, GoogLeNet, ShuffleNet, MobileNetv2, and DenseNet201 along with saliency methods Graph-Based Visual Saliency (GBVS), Cluster-based Saliency Detection (COS), and Spectral Residual (SPE). This work reported a maximum accuracy of 92.43% on the smaller dataset, while on the IP102 dataset, it was 61.93%. To categorize crop pests, Setiawan et al. [25] used the NBAIR, Xiel, and Xie2 insect datasets with 40, 24, and 40 classes, respectively. In their experiments on datasets, they used AlexNet, ResNet-50, ResNet-101, VGG-16, and VGG-19. For the insects' dataset mentioned earlier, the proposed CNN model achieved the highest classification accuracy of 96.75%, 97.47%, and 95.97%. To use convolutional neural networks for crop pest recognition in innate situations, Liu et al. [27] manually collected datasets of 10 pests (Gryllotalpa, Leafhopper, locust, Oriental fruit fly, Pieris rapae Linnaeus, Snail, Spodoptera Litura, Stinkbug, Cydia Pomonella, Weevil). Pre-trained models called VGG-16, VGG-19, ResNet50, ResNet152, and GoogLeNet were employed in this experiment. The accuracy of ResNet50 was 91.74%, ResNet152 was 92.9%, GoogleLeNet was 93.29%, VGG-16 was 91.44%, VGG-19 was 92.26%, and ResNet50 was 91.74%. Liu et al. [27] used the anchor-free region convolutional neural network (AF-RCNN) technique to detect agricultural pests in many categories. On a dataset with 24 classes of pests, this approach achieved 56.4% mean Average Precision (mAP) and 85.1% recall.

For an effective CNN-based pest localization and recognition, Li et al. [28] implemented data augmentation in the training phase, test time augmentation (TTA) approach, and region proposal network (RPN) techniques. The mAP for this model was 83.23%. To retrieve depth and spatial attention across many stages of the pyramid network, Liu et al. [29] manually collected dataset and implemented Global Activated Feature Pyramid Network (GaFPN). Next, a Locally Activated Region Proposal Network (LaRPN), an upgraded pest localization module, was put out to find the exact locations of the pest objects. In the end, a ResNet50 backbone was used, achieving an accuracy of 86.9%. Images were physically gathered from two greenhouse locations in Belgium by Nieuwenhuizen et al. [30]. After that, yellow sticky traps were utilized for insect detection and counting using Faster R-CNN with ResNet-v2 as its foundation. This approach had an accuracy rate of 87.4%. For large-scale multiclass pest detection and classification, Wang et al. [31] used the PestNet technique, which comprises three phases: pest feature extraction (a CNN backbone), pest areas search, and pest prediction (fuse RPN and PSSM). They achieved 75.46% mAP. The transfer learning (AlexNet) model was implemented by Dawei et al. [32]. The model's accuracy in identifying pests was 93.84%. Further, a pest dataset was created by Xia et al. [33] using manually gathered photos from search engines like Baidu and Google. The authors combined VGG19 and RPN models and obtained 89.22% insect detection and classification accuracy. Li et al. [34] used the transfer learning (DenseNet169) technique to classify pests in tomato plants. The collection includes 859 photos of tomato pests divided into 10 classifications, and 88.83% accuracy was attained by DenseNet-169. The IP102 dataset was reorganized by Li et al. [34] and given the name IP_RicePests. VGGNet, ResNet, and MobileNet networks were used to train the model. The experiment results demonstrate that all three classification networks, when paired with transfer learning, have good recognition precision, with the IP_RicePests dataset providing the best classification accuracy to careful adjustment of the ResNet50 network's parameters. ResNet50 had an accuracy of 87.41%, MobileNet had an accuracy of 86.44%, and VGG16 had an accuracy of 88.68%.

To recognize nine different types of diseases and pests on tomato plants, Sabanci et al. [35] integrated R-CNN, faster R-CNN, and SSD deep learning meta-learning with a visual geometry group network (VGG) and residual network [36]. A quick, precise, fine-grained object recognition model based on the YOLOv4 deep neural network was proposed by Roy et al. [37]. The proposed model's detection rate and mAP were 70.19 FPS and 96.29%, respectively. For the categorization of pests and diseases, Liu et al. [38] devised a self-supervised transformer-based pre-training technique employing Feature Relationship Conditional Filtering (FRCF) and Latent Semantic Masking Auto-Encoder (LSMAE). The accuracy rates for this study's utilization of the IP102, CPB, and Plant Village datasets were 74.69%, 76.99%, and 99.93%, respectively. Zhang et al. [40] employed the IP102 dataset to identify pests. The inceptionv3 model was used in this investigation, and the accuracy was 67.88%. On six classes of the IP102 dataset, Deepika and Arthi [39] constructed the Improved Mask Faster Region-Based Convolutional Neural Network (IMFR-CNN) model. The author has attained 99.2% accuracy in this investigation. Further, 20 classes from the IP102 dataset were used by Zhang et al. [40] for pest recognition. This research used the 97.8% accuracy on the Faster and Extensible Vision Transformer (FE-VIT) model. The Improved YOLO-X model was employed by Huang et al. [41] to spot forest pests. A precision of 53.6% was reached in this investigation using the IP102 dataset. Li et al. [42] implemented the Mask-RCNN ResNet50, Faster-RCNN ResNet101, and Yolov5 Darknet53 models for pest recognition. These methods each reached accuracy levels of 99.6%, 99.4%, and 97.6%. ResNeXt-50 (32 4d) model was used by Sanghavi et al. [43] to classify pest images using a residual neural network based on transfer learning using the IP102 dataset. This study had an accuracy rate of 86.90%. A manually compiled dataset was produced by Vishakha et al. [44] for the recognition and classification of crop pests using transfer learning. This study used the hunger games search-based deep convolutional neural network (HGS-DCNN) model with 99% accuracy. When used in intricate settings to spot numerous plant diseases, the model yields effective and efficient outcomes. For detecting pests, several researchers have investigated object identification methods based on DL [45], [46]. But none of this research covered the topic of identifying scale insect to preserve beneficial insects. Pest control is still not done in a way that is successful. Existing approaches lack to recognize huge classes of crop pests and unable to execute well with the presence of several image distortions like noise, blurring, color and light variations, etc. Therefore, there remains a need for a more effective approach to overcome the issues of existing works.

III. METHODOLOGY

This work proposed a model called the Faster-PestNet for the localization and division of numerous crop pests. Hence, we have altered the conventional Faster-RCNN approach by utilizing MobileNet as the base network, tuned on the pest samples to recognize the crop pests of various categories, and given the name Faster-PestNet. Real-time object identification on mobile devices is a good fit for MobileNet, an efficient CNN approach designed for mobile devices with a smaller footprint than conventional CNNs. Therefore, initially, the MobileNet is applied for calculating a distinctive set of image characteristics, later optimized and divided by the 2-step locator of the improved Faster-RCNN model and took the following actions:

IP 102 dataset, which contains the images of pests that belong to 102 classes, is used, and the local collected crops dataset is also used. We have used an annotated dataset to train a Faster-PestNet model using MobileNet as its foundation. The annotated photos are fed into the model, and the parameters are changed to reduce the discrepancy between the anticipated and actual bounding boxes.

After training the model, we used it to detect pests in new images by inputting them into the model and applying a threshold to the predicted bounding boxes to remove false positives.

The flow of the presented work is defined in Figure 1.

A. FASTER R-CNN

The object detection algorithm Faster R-CNN expands on the fundamental design of R-CNN and Fast R-CNN. A Region Proposal Network (RPN) and a Fast R-CNN detector comprise its primary parts. An image is sent into the RPN, a fully convolutional network, and outputs a list of object recommendations, each reflected by a bounding box and an objectness score.

The RPN can recognize things of various sizes since it operates across an image pyramid of various proportions. The RPN is trained to maximize two loss functions: a regression loss for forecasting the bounding box's coordinates and a classification loss for forecasting the objectness score of each proposal.

A proposal comprises an object or not is indicated by the objectness score, a binary classification score. This is how it is explained:

$$pi = \frac{1}{1 + e^{-wTxi}} \tag{1}$$

The feature vector of the proposal is represented by xi, and the weight vector of the classification layer is i and w.

The RPN generates a collection of 'k' proposals for every image and sends them to the Fast R-CNN detector for additional processing. The Fast R-CNN detector is a region-based detector that uses a Region of Interest (RoI) pooling layer to extract features from the RPN's proposal outputs. A fully connected network for classification and regression can be used once the RoI pooling layer has generated fixed-size feature maps for every set of rectangular RoIs.

B. FASTER-PestNet

The backbone network in the Faster R-CNN algorithm plays a critical role in extracting features from input images, which the RPN and classifier then utilize to detect objects. In this

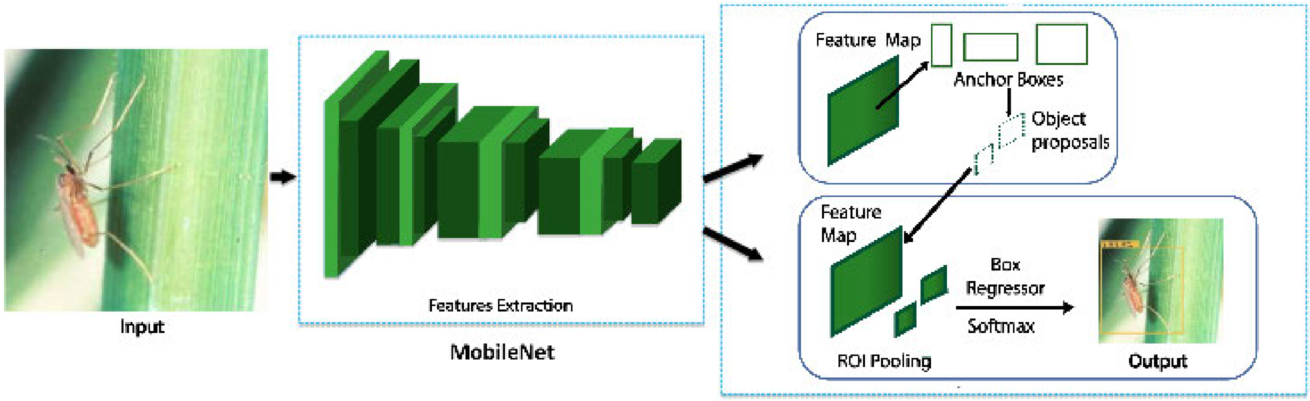

FIGURE 1. Architectural explanation of the Faster-PestNet technique.

project, we have utilized MobileNet as the backbone of the architecture to achieve this goal. The lightweight CNN architecture, MobileNet, is frequently utilized as the foundation for object detection algorithms due to its great efficiency and accuracy.

In this study, the ResNet backbone networks are swapped out for MobileNet backbone networks to serve as the Faster R-CNN algorithm's backbone. The major intuition for altering the base network of the conventional Faster-RCNN approach is that the ResNet approach is computationally more complex and unable to tackle the complicated sample transformational changes. To overwhelm the problems of the traditional model, we have employed the MobileNet approach as its feature extractor. The depth-wise separable convolutions used in the MobileNet design considerably lessen the number of specifications in the network while retaining good precision.

The Faster-PestNet algorithm's MobileNet backbone network can be characterized as follows:

1) MOBILENET BACKBONE

The MobileNet backbone network is made up of a number of fully connected (FC) layers after numerous depth-wise separable convolution layers. A size-related image acts as the input to the MobileNet backbone network, as given in Equation 2.

$$ImageSize = H \times W \times 3 \tag{2}$$

where the number of color channels is 3, and the image's height and width are H and W, respectively.

2) DEPTH-WISE SEPARABLE CONVOLUTION

There are two components to the depth-wise separable convolution layer: (i) depth-wise convolution and (ii) pointwise convolution. A 1 × 1 convolution is applied by pointwise convolution to aggregate the output of the depth-wise convolution, while the depth-wise convolution applies a single convolution filter to each input channel autonomously.

3) FULLY CONNECTED LAYERS

The output from depthwise separable convolution layers is mapped to a fixed-size feature map by the fully connected layers in the MobileNet backbone network, which the RPN and classifier can then use. The number of fully connected layers and their criteria can be adjusted based on the specific application.

C. OUTPUT FEATURE MAP

The MobileNet backbone network's output feature map is H/32 × W/32 × D, where D is the number of feature channels. It identifies objects in the image, and the RPN and classifier use this feature map as input.

D. RoI POOLING

The output feature map from the mobilenet backbone network is subjected to RoI pooling to take out a fixed-size feature vector for each region proposal after the RPN creates a series of region proposals. Each of the rectangular regions of the feature map corresponding to the region suggestions receives a max-pooling operation as part of the roi pooling operation.

E. FULLY CONNECTED LAYERS FOR CLASSIFICATION AND REGRESSION

After the RoI pooling layer's output, two distinct and completely connected layers for classification and regression are used. The object in the region proposal is classified using the classification layer, and the bounding box's coordinates are refined using the regression layer. The following can be used to represent the classification and regression layer equation:

$$fc_{cls} = ReLU(W_{cls} * h_{pool} + b_{cls}) \tag{3}$$

The weight matrix W_{cls}, the bias vector b_{cls}, the output of the RoI pooling layer h_{pool}, and the rectified linear unit activation function ReLU are all present. Equation 3 is for the classification layer, and Equation 4 is shown below for the regression layer:

$$fc_{reg} = W_{reg} * h_{pool} + b_{reg} \tag{4}$$

where W_{reg} is the weight matrix, and b_{reg} is the bias vector.

F. LOSS FUNCTION

The loss function is used to train the faster-pestnet algorithm. The faster-pestnet model is calculated from the output of the classification and regression layers. A Classification And Regression Loss Term Are Combined To Form The Loss Function. Equations 5 and 6 represent the classification and regression loss terms, respectively.

$$L_{cls}\left(p, p*\right) = -log\left(p*\right)\ if\ p* > 0\ else - log\left(1 - p*\right) \quad (5)$$

where the predicted probability of the object class is p, the ground truth probability of the object class is p*, and log is the natural logarithm.

$$L_{reg}(t, t*) = SmoothL1(t - t*) \quad (6)$$

where t is the predicted bounding box offset, t∗ is the ground truth bounding box offset, and SmoothL1 is the smooth L1 loss function defined as:

$$smoothL_1\left(x\right) = \begin{cases} 0.5x^2 & if\ |x| < 1 \\ |x| - 0.5 & otherwise \end{cases} \quad (7)$$

A succession of depth-wise separable convolution layers, fully connected layers, RoI pooling, and finally, distinct fully connected layers for classification and regression make up the MobileNet backbone of the Faster-PestNet method. The model is trained to precisely identify items in a picture using the loss function.

IV. EXPERIMENT DETAILS AND RESULTS

The details of execution and the evaluations performed to assess the results of the proposed approach are elaborated in this part. To thoroughly show the effectiveness of the Faster-PestNet model, we calculated pest recognition and division results via numerous experiments and correlated them with other models.

A. DATASET

For model tuning and testing, we have employed the IP102 dataset, which is a large-scale benchmark dataset for pest image classification and recognition tasks, consisting of 102 categories of crop pests commonly found in field areas. The IP102 dataset is challenging for picture categorization tasks. Dataset photos contain a broad range of perspective, scale, orientation, and illumination changes. The dataset has been utilized in several computer vision and machine learning research papers. Each of the 102 categories in the IP102 dataset has a different number of photos. More than 75,000 photos from 102 categories of pests are included in the dataset. The photos in the collection are RGB images with a resolution of 224 × 224 pixels. The pictures were gathered from online resources, including Google Images and Flickr. A training set, a validation set, and a test set are each divided into separate portions of the dataset. These collections each contain 56,846, 8,047, and 11,955 photos. Then, human specialists manually add object-level labels to the photos by hand, identifying one or more pests in the image. The IP102 dataset is a difficult benchmark for object recognition and detection tasks due to several distinctive features. These consist of:

- Diverse types of pests: The 102 categories in the IP102 dataset are diverse in nature.
- Occlusion and clutter: Many of the images in the IP102 dataset contain occlusions and clutter, such as multiple objects in the scene or objects partially obscured by other objects.
- Imbalanced class distribution: The number of images per category in the IP102 dataset varies broadly, with some categories having only a few images and others having thousands of images.

A few samples from the IP102 are given in Figure 2.

FIGURE 2. Samples from the IP102 dataset.

B. IMPLEMENTATION DETAILS

The Keras library is used in TensorFlow to implement the suggested framework. The Faster-PestNet model's final training parameters are detailed in Table 1. To create the final optimized model in our study, we varied the epochs, batch, and learning rate for the model's hyperparameters. The experiment used the Stochastic Gradient Descent (SGD) training optimizer amidst model learning rates of 0.0015. The epoch and batch size were 200 and 32. The input picture dimensions were set at 320 × 320, and the data were split into training authentication and test sets at random. 70% of the samples were utilized for training, 15% for validation, and 15% for examination.

C. EVALUATION PARAMETERS

We have employed a variety of quantitative indicators, including precision (P), recall (R), accuracy (Acc), and mAP, to assess the efficacy of the proposed approach. Following is how these metrics are computed:

$$P = TP/(TP + FP) \quad (8)$$

$$R = TP/(TP + FN) \quad (9)$$

$$Acc = (TP + TN)/(TP + TN + FP + FN) \quad (10)$$

True positive, true negative, false positive, and false negative situations are denoted by the letters TP, TN, FP, and FN. The pest in the image is regarded as TP if accurately identified; otherwise, it is regarded as FN. If the categorization is incorrect, the that is not visible in the photograph

TABLE 1. Training parameters for the proposed model.

Framework parameters	Value
Epochs	200
Learning rate	0.0015
Batch size	32

is categorized as TN; otherwise, it is classified as FP. The calculation for the mAP is displayed in Eq. (11), in which AP shows the average accuracy of every group, t is the examined image, and T denotes total test pictures.

$$mAP := \sum_{i=0}^{T} AP(t)/\mathrm{T} \tag{11}$$

D. PEST LOCALIZATION RESULTS

The accurate localization of pests is crucial for creating a successful automated pest recognition approach. So, we produced an experiment to evaluate how well the suggested framework can perform the localization of pests from the test samples.

We used all of the test photos from the IP102 database for the analysis, and the visual findings are given in Figure 3. We can deduct from the results that the suggested procedure Moreover, our method successfully finds pests despite challenging backdrop conditions, lighting, orientation shifts, and varied acquisition angles effectively finds pests of various sizes, shapes, and colors.

The suggested framework's capacity to localize by using key points estimate enables it to distinguish and detect pests of different types accurately and effectively. We calculated the mAP to quantitatively estimate the localization and recognition for a number of pest categories. Specifically, we have attained a mAP of 0.8243, indicating our approach's robustness in recognizing the pests from such a complex data sample due to its high recall ability.

E. PEST CLASSIFICATION RESULTS

It is crucial to accurately classify different pests to show that a model is reliable. Depending on the crop category, several different sorts of pests may be in a crop-growing region. In order to assess the effectiveness of the suggested procedure for categorizing pests based on numerous hierarchical crop groups, we experimented.

The trained Faster-PestNet framework is evaluated for each test picture from the IP102 data sample to complete this task. Figures 4, 5, 6, and 7 display the precision, recall, F1 score, and accuracy results for the suggested method in crop-based pest classification. The findings show that the proposed framework achieves precision, recall, and F1 scores, with 83%, 81%, 82%, and 82.43% accuracy for all crop-specific classes.

The efficacy of the applied key points computing approach, which accurately and reliably depicts each pest class, is the cause of the strong pest categorization presentation of the suggested model. Consequently, it is not wrong to say that Faster-PestNet executes well for crop-wise pest identification, proving the value of the suggested approach.

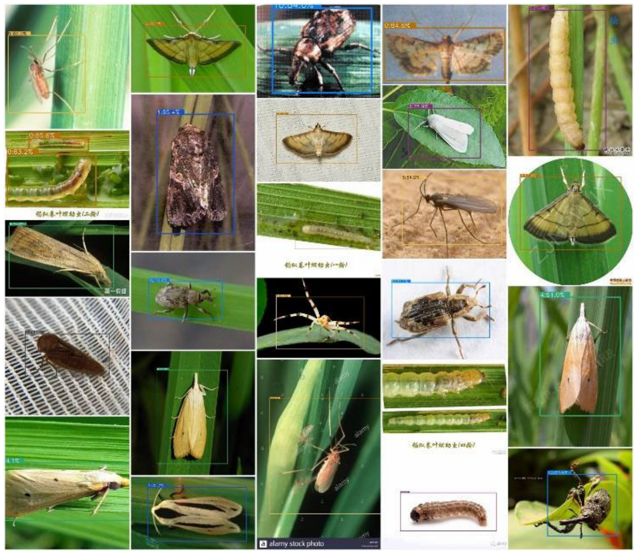

FIGURE 3. Localization visuals attained for the Faster-PestNet.

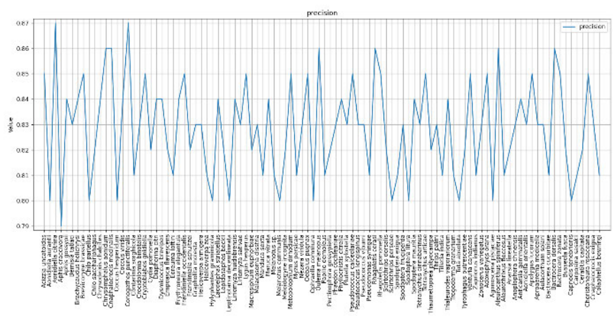

FIGURE 4. Precision values for the Faster-PestNet.

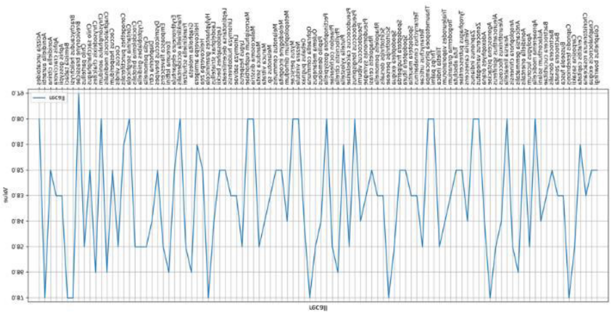

FIGURE 5. Recall values for the Faster-PestNet.

F. EVALUATION OF FASTER RCNN MOBILENET MODEL

An accurate and nominative set of features is required for effective target recognition. We analyzed to compare the proposed Faster-PestNet model to other deep feature extraction frameworks to compare pest detection and classification results against them. We evaluated our values with several base models, including Alexnet [47], GoogleNet [48],

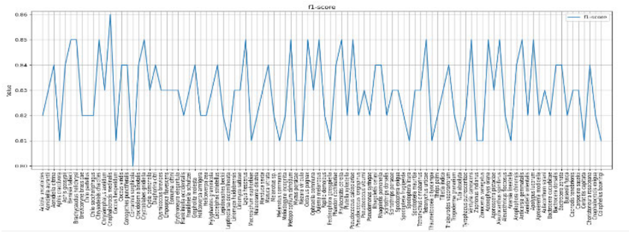

FIGURE 6. F1-score values for the Faster-PestNet.

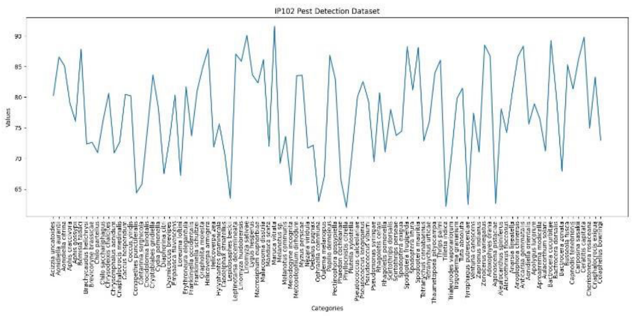

FIGURE 7. Accuracy values for the Faster-PestNet.

VGGNet [49], ResNet-50 [50], ResNet-101 [50], Inception V4 [51], HourGlass104 [52], EfficientNet [53]. Additionally, the network complexities of all frameworks are also discussed to check the efficiency of the approaches as well. The results, in terms of accuracy and the model parameters, are given in Table 2.

The values in Table 2 show that we perform better than all other DL frameworks, with the highest accuracy value of 82.43%. Moreover, the proposed approach contains fewer parameters than all other approaches, giving an edge to the Faster-PestNet in terms of model complexity. We acquire better results than all other DL models because of the model's capacity to extract the visual attributes of the input samples finely which enhances the recognition power of the Faster-PestNet model.

G. RESULTS ANALYSIS WITH OBJECT IDENTIFICATION METHODS

We equate the suggested model's conduct to various cutting-edge methods for identifying pest objects. Precise pest recognition is essential since a cluttered background could fool the predictor when the target is not immediately visible. The presence of numerous s could make detection quite challenging. Proper localization can enhance precision to a greater extent by minimizing unnecessary background information. We examined a number of one-stage object identification models, which have been demonstrated to be effective on the COCO dataset [54], including SSD [50], YOLOv3 [55], RefineDet [56], and CornerNet [57]. Other two-stage detectors, such as Fast R- CNN [18] and Faster R- CNN [58], were also examined. We tested these models' performance on the IP102 dataset to see how well they could localize pests in various challenging situations, such as

TABLE 2. Faster-PestNet performance comparison with other deep learning models.

Method	Accuracy
AlexNet	41.8
GoogleNet	43.5
VGGNet	48.2
ResNet-50	57.39
ResNet-101	53.18
Inception V4	47.8
HourGlass 104	54.63
EfficientNet	60.2
DenseNet-121	54.71
Faster-PestNet	82.43

when complicated environmental noise, brightness, hue, size, and contour variations were present. We calculated the mAP measure, a typical metric used in object identification tasks, to conduct the performance study. We have also generated test times to assess the computational complexity of each model. The performance of numerous object detection techniques with different backbones for identifying pests is compared in Table 3 in terms of mAP and inference time.

Table 3 shows that the suggested method for identifying pests outperforms the alternative. 2-step object identifiers like Fast R-CNN and Faster R-CNN perform worse. These methods are computationally expensive because they use anchor boxes to pinpoint the possible area of interest before performing division and regression to pick the appropriate box. In contrast, the 1-stage models RefineDet, SSD, and YOLO-v3 perform better since they determine the location and category of the item directly. However, when the initial applications of these technologies are evaluated in this work, they are found to be insufficient for locating and identifying pests in settings with drastic changes in lighting. Furthermore, it is shown that one-stage detectors compute more quickly than two-stage detector models.

Our model, Faster-PestNet, successfully overcomes the disadvantage of previous methodologies by utilizing MobileNet as its backbone network. The Mobilenet backbone enables the improved Faster RCN to learn more representative qualities such as noise, luminance, and diversity in color, size, and form. It allows for improved localization and classification of pests into distinct groups. We calculated the mAP measure, a typical metric used in object recognition tasks, to conduct the performance study. We have also generated test times to assess the computational complexity of each model. So, based on Table 3 findings, we can conclude that the suggested method for identifying pests outperforms the alternative in terms of pest recognition and inference time.

H. PERFORMANCE COMPARISON WITH EXISTING APPROACHES

The following part compares the classification performance of our technique with findings from other studies [25], [26], [27], [58], [59], [60], [61], [62], [63] using the same Dataset,

TABLE 3. Faster-PestNet comparison with other object locators.

Method	Base	mAP
Fast R-CNN	VGG-16	46.12
Fast R-CNN	DenseNet-100	47.34
Faster R-CNN	VGG-16	47.87
Faster R-CNN	DenseNet-100	48.12
Refinedet	VGG-16	49.01
Refinedet	DenseNet-100	50.25
SSD	VGG-16	47.21
SSD	DenseNet-100	49.17
YOLOv3	DarkNet-53	50.64
YOLOv3	DenseNet-100	50.95
Proposed (Faster-PestNet)	Mobilenet	82.43

TABLE 4. Faster-PestNet comparison with existing approaches.

Reference	Method	Accuracy (%)
[58]	InceptionNetV3	57.08
[59]	GAEnsemble	67.13
[60]	EquisiteNet	52.32
[61]	FR-ResNet-50	55.24
[27]	DFF-ResNet-82	55.43
[26]	FusionSum-Densenet201	61.93
[62]	Ensemble Model	74.13
[25]	MobileNetV2+Sparse+CutMix+ DynamicLR	71.32
[63]	PCNet	73.70
Proposed	Faster-PestNet	82.43

IP102. Based on their average accuracy, the results of the pest classification are contrasted with current techniques in Table 4.

When writers in [58] trained various deep learning models to classify pest species, they discovered that this approach had the highest overall average accuracy of 57.08% with InceptionNetV3. The author manually resized and augmented the dataset before training this network. Ayan et al. [59] combined CCNs with ensemble methods, particularly GAEnsemble, to improve the division presentation. Like [26], the writers merge CNN and the saliency approach to make an ensemble of predictors at the output layer using the fusion-sum methodology. However, the ensemble weights computation used by both approaches [26], [59] resulted in a poor calculating time that cost them an accuracy of 61.93% and 67.13%, respectively. The EquisiteNet model with double fusion, squeeze and excitation, and max-feature expansion blocks was employed by Zhou and Su [60]. The accuracy of the model was 52.32%. The methods in [27] and [61] included feature reuse and feature fusion procedures in addition to the modified Resnet block for efficient feature calculation. They were accurate to a respective degree of 55.24% and 55.43%. Using the IP102 dataset [62] also attempted several tests. ResNet-50, RAN (Residual Attention Network), FPN (Feature Pyramid Network), and MMAL-NET were all used in these investigations. The ensemble of all models was run, and the best results with a precision of 74.13% were obtained. The author [25] suggests an effective training method that takes advantage of CutMix augmentation, freezing layers, and sparse regularization to optimize small-sized MobileNetV2 models. The maximum accuracy, 71.32%, was achieved by amalgamating those techniques throughout training. An effective and lightweight pest classification model called PCNet was proposed in [63]. To improve the depiction of pest keypoints in complicated and comparable backgrounds and pay closer heed to pest details, PCNet is an improved approach based on EfficientNet v2 that integrates the attention mechanism. Additionally, PCNet takes advantage of the keypoints fusion unit to minimize the loss of pest information when multiple depth layers are down-sampled.

FIGURE 8. Faster-PestNet visual results on the local crops dataset.

These data clearly show that the proposed Faster-PestNet outperforms past studies, with an average accuracy of 82.43%. The result from previous layers as input to all following layers effectively computes feature maps with MobileNet's improved performance results from linking. The Faster-PestNet architecture locates and classifies the pests using calculated features. Consequently, the proposed version performs much finer for challenging Dataset IP102 regarding pest recognition and classification. In addition, our technique is computationally competent and trustworthy enough to recognize s with better precision than others. As a result, we may conclude that our method has enormous promise for using drones to classify target pests in the field.

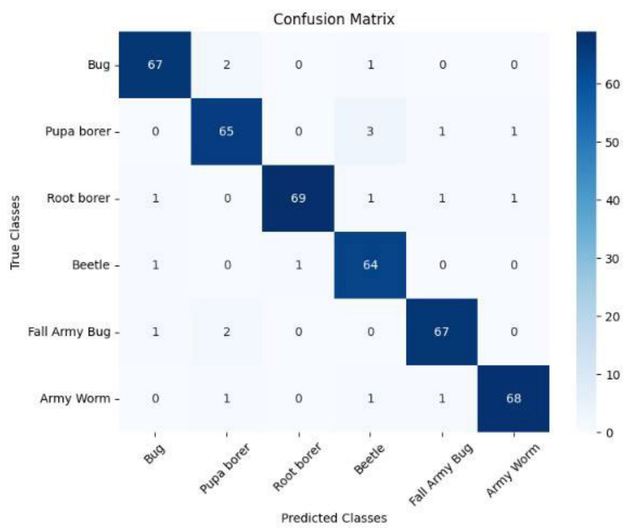

FIGURE 9. Faster-PestNet confusion matrix on the local crops dataset.

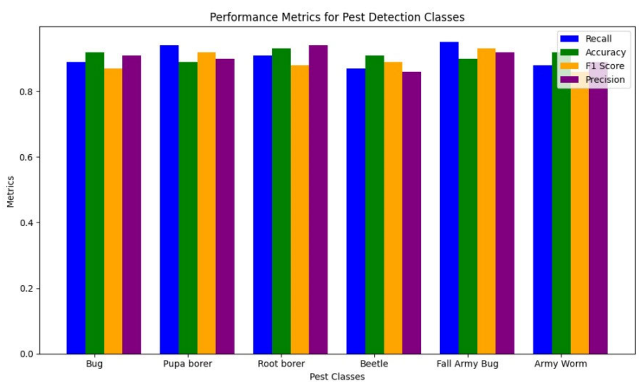

FIGURE 10. Faster-PestNet performance analysis on the local crops dataset.

I. GENERALIZATION ABILITY TESTING

To further show the robustness of the proposed Faster-PestNet approach. We have collected a local crops dataset comprising a total of 1950 images from various site areas and labeled them into six classes with the help of domain experts. These classes are named 'Bug', 'Pupa Borer', 'Root Borer', 'Beetle', 'Fall Army Bug', and 'Army Worm'.

We tuned the Faster-PestNet on this data sample using the 80-20% division mechanism for model learning and evaluation. Figure 8 shows visuals attained and clearly shows that our approach can diagnose the Pest of numerous types from real-world examples under complicated background settings.

Further, the confusion matrix for this locally crops dataset is reported in Figure 9 to show the recognition ability of our approach. The values in Figure 9 prove that the Faster-PestNet approach is proficient in recognizing all classes of pests with a high recall rate. We have also measured other performance measures like precision, recall, F1, and accuracy metrics, and attained values are given in Figure 10. We have obtained 95.24%, 95.26%, 95.23%, and 95.24% for this local crops dataset's precision, recall, F1-score, and accuracy metrics. Based on the conducted analysis, it can be concluded that the presented Faster-PestNet approach is proficient in recognizing all classes of pests due to its high generalization power and capability to overwhelm the model over-fitting problem.

V. CONCLUSION

In this study, we have provided a cost-effective DL system for the automated spotting and division of crop pests. Specifically, a model named the Faster-PestNet is proposed in which the MobileNet approach is used as a core network for collecting dense features. We tested our method using the IP102 dataset collection, representing an extensive and challenging standard collection for pest identification made up of in-field collected photos. We have demonstrated the viability of our method for practical pest surveillance tasks through considerable experiments. The results indicated that our system could reliably localize and categorize pests of various types, even in the context of complicated backgrounds and fluctuations in different pest forms, hues, dimensions, positions, and luminance. Further, we have a local crop dataset and evaluated our approach to show better generalization capability. All reported numeric and pictorial evaluations have proved our approach's effectiveness in recognizing huge pests. As a future consideration, we are willing to design a further enhanced DL approach to improve the classification results by taking into account other strategies like feature fusion, etc. Further, we are motivated to evaluate the proposed technique in other agriculture-related applications, such as recognizing the crop diseases caused by pests.

REFERENCES

[1] J. Bruinsma, "The resource outlook to 2050: By how much do land, water and crop yields need to increase by 2050," in *Proc. Exp. Meeting Feed World*, 2009, pp. 24–26.

[2] S. Neethirajan, "The role of sensors, big data and machine learning in modern animal farming," *Sens. Bio-Sens. Res.*, vol. 29, Aug. 2020, Art. no. 100367.

[3] T. Bai, J. Lin, G. Li, H. Wang, P. Ran, Z. Li, D. Li, Y. Pang, W. Wu, and G. Jeon, "A lightweight method of data encryption in BANs using electrocardiogram signal," *Future Gener. Comput. Syst.*, vol. 92, pp. 800–811, Mar. 2019.

[4] J. Qin, F. Liu, K. Liu, G. Jeon, and X. Yang, "Lightweight hierarchical residual feature fusion network for single-image super-resolution," *Neurocomputing*, vol. 478, pp. 104–123, Mar. 2022.

[5] M. T. Mallick, S. Biswas, A. K. Das, H. N. Saha, A. Chakrabarti, and N. Deb, "Deep learning based automated disease detection and pest classification in Indian *mung bean*," *Multimedia Tools Appl.*, vol. 82, no. 8, pp. 12017–12041, Mar. 2023.

[6] K. Rimal, K. B. Shah, and A. K. Jha, "Advanced multi-class deep learning convolution neural network approach for insect pest classification using TensorFlow," *Int. J. Environ. Sci. Technol.*, vol. 20, no. 4, pp. 4003–4016, Apr. 2023.

[7] W. Albattah, M. Masood, A. Javed, M. Nawaz, and S. Albahli, "Custom CornerNet: A drone-based improved deep learning technique for large-scale multiclass pest localization and classification," *Complex Intell. Syst.*, vol. 9, no. 2, pp. 1299–1316, Apr. 2023.

[8] G. Pajares, "Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs)," *Photogramm. Eng. Remote Sens.*, vol. 81, no. 4, pp. 281–329, 2015.

[9] E. F. I. Raj, M. Appadurai, and K. Athiappan, "Precision farming in modern agriculture," in *Smart Agriculture Automation Using Advanced Technologies: Data Analytics and Machine Learning, Cloud Architecture Automation and IoT*. Singapore: Springer, 2022, pp. 61–87.

[10] A. Bulat and G. Tzimiropoulos, "Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 109–117.

[11] P. Zhou, B. Ni, C. Geng, J. Hu, and Y. Xu, "Scale-transferrable object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 528–537.

[12] T. Roska and L. O. Chua, "The CNN universal machine: An analogic array computer," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, no. 3, pp. 163–173, Mar. 1993.

[13] L. R. Medsker and L. Jain, "Recurrent neural networks," *Des. Appl.*, vol. 5, p. 2, Dec. 2001.

[14] A. Kamilaris and F. X. Prenafeta-Boldu, "Deep learning in agriculture: A survey," *Comput. Electron. Agricult.*, vol. 147, pp. 70–90, Apr. 2018.

[15] P. Sermanet, A. Frome, and E. Real, "Attention for fine-grained categorization," 2014, *arXiv:1412.7054*.

[16] H. Kataoka, K. Iwata, and Y. Satoh, "Feature evaluation of deep convolutional neural networks for object recognition and detection," 2015, *arXiv:1509.07627*.

[17] L. Wang, S. Guo, W. Huang, and Y. Qiao, "Places205-VGGNet models for scene recognition," 2015, *arXiv:1508.01667*.

[18] S. Targ, D. Almeida, and K. Lyman, "ResNet in ResNet: Generalizing residual architectures," 2016, *arXiv:1603.08029*.

[19] A. Raj, V. P. Namboodiri, and T. Tuytelaars, "Subspace alignment based domain adaptation for RCNN detector," 2015, *arXiv:1507.05578*.

[20] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.

[21] X. Zhao, W. Li, Y. Zhang, T. A. Gulliver, S. Chang, and Z. Feng, "A faster RCNN-based pedestrian detection system," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2016, pp. 1–5.

[22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[23] J. G. A. Barbedo, "Detecting and classifying pests in crops using proximal images and machine learning: A review," *AI*, vol. 1, no. 2, pp. 312–328, Jun. 2020.

[24] A. Kumar, Z. J. Zhang, and H. Lyu, "Object detection in real time based on improved single shot multi-box detector algorithm," *EURASIP J. Wireless Commun. Netw.*, vol. 2020, no. 1, pp. 1–18, Dec. 2020.

[25] A. Setiawan, N. Yudistira, and R. C. Wihandika, "Large scale pest classification using efficient convolutional neural network with augmentation and regularizers," *Comput. Electron. Agricult.*, vol. 200, Sep. 2022, Art. no. 107204.

[26] L. Nanni, G. Maguolo, and F. Pancino, "Insect pest image detection and recognition based on bio-inspired methods," *Ecological Informat.*, vol. 57, May 2020, Art. no. 101089.

[27] W. Liu, G. Wu, F. Ren, and X. Kang, "DFF-ResNet: An insect pest recognition model based on residual networks," *Big Data Mining Anal.*, vol. 3, no. 4, pp. 300–310, Dec. 2020.

[28] R. Li, X. Jia, M. Hu, M. Zhou, D. Li, W. Liu, R. Wang, J. Zhang, C. Xie, L. Liu, F. Wang, H. Chen, T. Chen, and H. Hu, "An effective data augmentation strategy for CNN-based pest localization and recognition in the field," *IEEE Access*, vol. 7, pp. 160274–160283, 2019.

[29] L. Liu, C. Xie, R. Wang, P. Yang, S. Sudirman, J. Zhang, R. Li, and F. Wang, "Deep learning based automatic multiclass wild pest monitoring approach using hybrid global and local activated features," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7589–7598, Nov. 2021.

[30] A. Nieuwenhuizen, J. Hemming, and H. Suh, "Detection and classification of insects on stick-traps in a tomato crop using faster R-CNN," Agro Food Robot., Wageningen Univ. Res., Wageningen, The Netherlands, Tech. Rep., 2018.

[31] L. Liu, R. Wang, C. Xie, P. Yang, F. Wang, S. Sudirman, and W. Liu, "PestNet: An end-to-end deep learning approach for large-scale multi-class pest detection and classification," *IEEE Access*, vol. 7, pp. 45301–45312, 2019.

[32] W. Dawei, D. Limiao, N. Jiangong, G. Jiyue, Z. Hongfei, and H. Zhongzhi, "Recognition pest by image-based transfer learning," *J. Sci. Food Agricult.*, vol. 99, no. 10, pp. 4524–4531, Mar. 2019.

[33] D. Xia, P. Chen, B. Wang, J. Zhang, and C. Xie, "Insect detection and classification based on an improved convolutional neural network," *Sensors*, vol. 18, no. 12, p. 4169, Nov. 2018.

[34] Z. Li, X. Jiang, X. Jia, X. Duan, Y. Wang, and J. Mu, "Classification method of significant rice pests based on deep learning," *Agronomy*, vol. 12, no. 9, p. 2096, Sep. 2022.

[35] K. Sabanci, M. F. Aslan, E. Ropelewska, M. F. Unlersen, and A. Durdu, "A novel convolutional-recurrent hybrid network for sunn pest-damaged wheat grain detection," *Food Anal. Methods*, vol. 15, no. 6, pp. 1748–1760, Jun. 2022.

[36] A. Fuentes, S. Yoon, S. Kim, and D. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, p. 2022, Sep. 2017.

[37] A. M. Roy, R. Bose, and J. Bhaduri, "A fast accurate fine-grain object detection model based on YOLOv4 deep neural network," *Neural Comput. Appl.*, vol. 34, pp. 3895–3921, Jan. 2022.

[38] H. Liu, Y. Zhan, H. Xia, Q. Mao, and Y. Tan, "Self-supervised transformer-based pre-training method using latent semantic masking auto-encoder for pest and disease classification," *Comput. Electron. Agricult.*, vol. 203, Dec. 2022, Art. no. 107448.

[39] P. Deepika and B. Arthi, "Prediction of plant pest detection using improved mask FRCNN in cloud environment," *Meas., Sensors*, vol. 24, Dec. 2022, Art. no. 100549.

[40] L. Zhang, J. Du, and R. Wang, "FE-VIT: A faster and extensible vision transformer based on self pre-training for pest recognition," *Proc. SPIE*, vol. 12349, pp. 35–42, Oct. 2022.

[41] J. Huang, Y. Huang, H. Huang, W. Zhu, J. Zhang, and X. Zhou, "An improved YOLOX algorithm for forest pest detection," *Comput. Intell. Neurosci.*, vol. 2022, Aug. 2022, Art. no. 5787554.

[42] W. Li, T. Zhu, X. Li, J. Dong, and J. Liu, "Recommending advanced deep learning models for efficient insect pest detection," *Agriculture*, vol. 12, no. 7, p. 1065, Jul. 2022.

[43] C. Li, T. Zhen, and Z. Li, "Image classification of pests with residual neural network based on transfer learning," *Appl. Sci.*, vol. 12, no. 9, p. 4356, Apr. 2022.

[44] V. B. Sanghavi, H. Bhadka, and V. J. Dubey, "Hunger games search based deep convolutional neural network for crop pest identification and classification with transfer learning," *Evolving Syst.*, vol. 14, pp. 649–671, Jul. 2022.

[45] W. Xia, D. Han, D. Li, Z. Wu, B. Han, and J. Wang, "An ensemble learning integration of multiple CNN with improved vision transformer models for pest classification," *Ann. Appl. Biol.*, vol. 182, no. 2, pp. 144–158, Mar. 2023.

[46] S. Wang, Q. Zeng, W. Ni, C. Cheng, and Y. Wang, "ODP-transformer: Interpretation of pest classification results using image caption generation techniques," *Comput. Electron. Agricult.*, vol. 209, Jun. 2023, Art. no. 107863.

[47] A. Krizhevsky and V. J. Nair, "Cifar-100 (Canadian Institute for advanced research). 30 [65] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, p. 26.

[48] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[49] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[50] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision—ECCV*. Amsterdam, The Netherlands: Springer, 2016, pp. 21–37.

[51] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Improved inception-residual convolutional neural network for object recognition," *Neural Comput. Appl.*, vol. 32, no. 1, pp. 279–293, Jan. 2020.

[52] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Computer Vision—ECCV*. Amsterdam, The Netherlands: Springer, 2016, pp. 483–499.

[53] Ü. Atila, M. Uçar, K. Akyol, and E. J. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," *Ecol. Inform.*, vol. 61, Mar. 2021, Art. no. 101182.

[54] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Computer Vision—ECCV*. Zurich, Switzerland: Springer, 2014, pp. 740–755.

[55] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[56] S. Zhang, L. Wen, X. Bian, Z. Lei, and S. Z. Li, "Single-shot refinement neural network for object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4203–4212.

[57] H. Law and J. Deng, "CornerNet: Detecting objects as paired keypoints," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 734–750.

[58] M. T. Reza, N. Mehedi, N. A. Tasneem, and M. A. Alam, "Identification of crop consuming insect pest from visual imagery using transfer learning and data augmentation on deep neural network," in *Proc. 22nd Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2019, pp. 1–6.

[59] E. Ayan, H. Erbay, and F. Varçin, "Crop pest classification with a genetic algorithm-based weighted ensemble of deep convolutional neural networks," *Comput. Electron. Agricult.*, vol. 179, Dec. 2020, Art. no. 105809.

[60] S.-Y. Zhou and C.-Y. Su, "Efficient convolutional neural network for pest recognition—ExquisiteNet," in *Proc. IEEE Eurasia Conf. IoT, Commun. Eng. (ECICE)*, Oct. 2020, pp. 216–219.

[61] F. Ren, W. Liu, and G. Wu, "Feature reuse residual networks for insect pest recognition," *IEEE Access*, vol. 7, pp. 122758–122768, 2019.

[62] H. T. Ung, H. Q. Ung, and B. T. Nguyen, "An efficient insect pest classification using multiple convolutional neural network based models," 2021, *arXiv:2107.12189*.

[63] T. Zheng, X. Yang, J. Lv, M. Li, S. Wang, and W. Li, "An efficient mobile model for insect image classification in the field pest management," *Eng. Sci. Technol., Int. J.*, vol. 39, Mar. 2023, Art. no. 101335.

[64] X. Wu, C. Zhan, Y.-K. Lai, M.-M. Cheng, and J. Yang, "IP102: A large-scale benchmark dataset for insect pest recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8779–8788.

FAROOQ ALI received the B.Sc. degree in software engineering from the University of Engineering and Technology at Taxila, and the M.Sc. degree in software engineering from Riphah International University Islamabad, with a focus on software engineering discipline. He is currently pursuing the Ph.D. degree with UET Taxila. His current research interests include image processing, medical image analysis, and computer vision.

HUMA QAYYUM received the M.Sc. degree in software engineering from the NUST College of Electrical and Mechanical Engineering, Pakistan, in 2009, and the Ph.D. degree in software engineering from the University of Engineering and Technology at Taxila, Pakistan, in2019. Her current research interests include image processing, medical image analysis, and computer vision.

MUHAMMAD JAVED IQBAL received the M.Sc. degree in computer science from the University of Agriculture, Faisalabad, Pakistan, in 2001, the M.S./M.Phil. degree in computer science from International Islamic University Islamabad, Pakistan, in 2008, and the Ph.D. degree in computer science/information technology from Universiti Teknologi PETRONAS, Malaysia, in February 2015. He is currently a HEC-approved Ph.D. Supervisor and an Assistant Professor with the Computer Science Department, University of Engineering and Technology, Taxila, Pakistan. After the Ph.D. studies, he has been actively involved in research. He has more than 20 international publications which include four ISI-indexed impact factor journals.